

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/EP 94/01625

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 5 C12N15/31 C12N9/80 C12Q1/68 C12P21/08 A61K39/106
 G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 5 C12N C12Q C12P A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GASTROENTEROLOGY, vol.104, no.4, April 1993, ELSEVIER, NEW YORK, U.S.; page A699 R.L. FERRERO ET AL. 'Molecular evidence demonstrating significant homology between the urease polypeptides of Helicobacter felis and Helicobacter pylori' Digestive disease week and the 94th annual meeting of the american gastroenterological association, May 15-21, 1993; Boston, Massachusetts, US; * page A699, left column, paragraph 2 * --- -/-	10,11,20
Y		7-9, 12-16, 22, 24-28, 31-35



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

1

Date of the actual completion of the international search

10 October 1994

Date of mailing of the international search report

27.10.94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl.
Fax (+ 31-70) 340-3016

Authorized officer

Hornig, H

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/EP 94/01625

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,93 07273 (INSTITUT PASTEUR) 15 April 1993 cited in the application * the whole document * ---	7-16, 22, 24-28, 31-35
X	INFECTION AND IMMUNITY, vol.60, no.5, May 1992, AM. SOC. MICROBIOL., BALTIMORE, US; pages 1946 - 1951 B.E. DUNN ET AL. 'Identification and purification of a cpn60 heat shock protein homolog from Helicobacter pylori' cited in the application the whole document ---	17
X	INFECTION AND IMMUNITY, vol.60, no.5, May 1992, AM. SOC. MICROBIOL., BALTIMORE, US; pages 2125 - 2127 D.J. EVANS ET AL. 'Urease-associated heat shock protein of Helicobacter pylori' cited in the application the whole document ---	17
Y	WO,A,90 04030 (INSTITUT PASTEUR) 19 April 1990 the whole document ---	10-16, 20,21, 24,34,35
Y	WO,A,91 09049 (RESEARCH EXPLOITATION LIMITED) 27 June 1991 the whole document ---	10-16, 20,21, 24,34,35
Y	J. CLIN. MICROBIOL., vol.30, no.3, March 1992, AM. SOC. MICROBIOL., WASHINGTON, DC,US; pages 739 - 741 P.A. FOXALL ET AL. 'Use of polymerase chain reaction-amplified Helicobacter pylori urease structural genes for differentiation of isolates' the whole document ---	10-16, 20,21, 24,34,35
P,X	WO,A,94 06474 (GALAGEN INC.) 31 March 1994 the whole document ---	27-32, 37,38
P,X	WO,A,93 18150 (BIOCINE-SCLAVO S.P.A.) 16 September 1993 the whole document ---	17,20, 22,24,25 34,35
1		-/-

INTERNATIONAL SEARCH REPORT

Internal Application No	PCT/EP 94/01625
-------------------------	-----------------

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>MOLEC. MICROBIOL., vol.9, no.2, 14 July 1993, BLACKWELL SCI. PUB., OXFORD, UK; pages 323 - 333 R.L. FERRERO AND A. LABIGNE 'Cloning, expression and sequencing of Helicobacter felis urease genes' see page 324, left column, paragraph 2 - page 326, right column, paragraph 1; figures 1,3 ----</p>	7,8, 10-12,16
P,X	<p>ABSTR. GEN. MEET. AM. SOC. MICROBIOL., vol.93, no.0, 19 May 1993 page 127 S. SUERBAUM AND A. LABIGNE 'Cloning and sequencing of the HSPA and HSPB heat shock protein encoding genes of Helicobacter pylori' 93rd general meeting of the american society for microbiology, Atlanta, Georgia, USA, May 16-20, 1993; abstract no. D-182; see abstract -----</p>	17-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internal Application No

PCT/EP 94/01625

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9307273	15-04-93	FR-A-	2682122	09-04-93
		CA-A-	2120527	15-04-93
		EP-A-	0610322	17-08-94
-----	-----	-----	-----	-----
WO-A-9004030	19-04-90	FR-A-	2637612	13-04-90
		EP-A-	0367644	09-05-90
		JP-T-	3501928	09-05-91
-----	-----	-----	-----	-----
WO-A-9109049	27-06-91	NONE		-----
-----	-----	-----	-----	-----
WO-A-9406474	31-03-94	AU-B-	4924893	12-04-94
-----	-----	-----	-----	-----
WO-A-9318150	16-09-93	NONE		-----
-----	-----	-----	-----	-----

(vi) ORIGINAL SOURCE:

(A) ORGANISM : H. felis

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..591

(D) OTHER INFORMATION: /standard_name= "URE I"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATG TTA GGT CTT GTG TTA TTG TAT GTT GCG GTC GTG CTG ATC AGC AAC Met Leu Gly Leu Val Leu Leu Tyr Val Ala Val Val Leu Ile Ser Asn 1 5 10 15	48
GGA GTT ACT GGG CTT GCA AAT GTG GAT GCC AAA AGC AAA GCC ATC ATG Gly Val Ser Gly Leu Ala Asn Val Asp Ala Lys Ser Lys Ala Ile Met 20 25 30	96
AAC TAC TTT GTG GGG GGG GAC TCT CCA TTG TGT GTA ATG TGG TCG CTA Asn Tyr Phe Val Gly Gly Asp Ser Pro Leu Cys Val Met Trp Ser Leu 35 40 45	144
TCA TCT TAT TCC ACT TTC CAC CCC ACC CCC CCT GCA ACT GGT CCA GAA Ser Ser Tyr Ser Thr Phe His Pro Thr Pro Pro Ala Thr Gly Pro Glu 50 55 60	192
GAT GTC GCG CAG GTG TCT CAA CAC CTC ATT AAC TTC TAT GGT CCA GCG Asp Val Ala Gln Val Ser Gln His Leu Ile Asn Phe Tyr Gly Pro Ala 65 70 75 80	240
ACT GGT CTA TTG TTT GGT TTT ACC TAC TTG TAT GCT GCC ATC AAC AAC Thr Gly Leu Leu Phe Gly Phe Thr Tyr Leu Tyr Ala Ala Ile Asn Asn 85 90 95	288
ACT TTC AAT CTC GAT TGG AAA CCC TAT GGC TGG TAT TGC TTG TTT GTA Thr Phe Asn Leu Asp Trp Lys Pro Tyr Gly Trp Tyr Cys Leu Phe Val 100 105 110	336
ACC ATC AAC ACT ATC CCA GCG GCC ATT CTT TCT CAC TAT TCC GAT GCG Thr Ile Asn Thr Ile Pro Ala Ala Ile Leu Ser His Tyr Ser Asp Ala 115 120 125	384
CTT GAT GAT CAC CGC CTC TTA GGA ATC ACT GAG GGC GAT TGG TGG GCT Leu Asp Asp His Arg Leu Leu Gly Ile Thr Glu Gly Asp Trp Trp Ala 130 135 140	432
TTC ATT TGG CTT GCT TGG GGT GTT TTG TGG CTC ACT GGT TGG ATT GAA Phe Ile Trp Leu Ala Trp Gly Val Leu Trp Leu Thr Gly Trp Ile Glu 145 150 155 160	480
TGC GCA CTT GGT AAG AGT CTA GGT AAA TTT GTT CCA TGG CTT GCC ATC Cys Ala Leu Gly Lys Ser Leu Gly Lys Phe Val Pro Trp Leu Ala Ile 165 170 175	528

98

GTC GAG GGC GTG ATC ACC GCT TGG ATT CCT GCT TGG CTA CTC TTT ATC
 Val Glu Gly Val Ile Thr Ala Trp Ile Pro Ala Trp Leu Leu Phe Ile
 180 185 190

CAA CAC TGG TCT TGA
 591
 Gln His Trp Ser
 195

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 199 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM : H. felis
- (v) SEQUENCE DESCRIPTION: SEQ ID

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Lys Gly Trp Met Leu Gly Leu Val Leu Leu Tyr Val Ala Val Val Leu
 1 5 10 15

Ile Ser Asn Gly Val Ser Gly Leu Ala Asn Val Asp Ala Lys Ser Lys
20 25 30

Ala Ile Met Asn Tyr Phe Val Gly Gly Asp Ser Pro Leu Cys Val Met
35 40 45

Trp Ser Leu Ser Ser Tyr Ser Thr Phe His Pro Thr Pro Pro Ala Thr
50 55 60

Gly Pro Glu Asp Val Ala Gln Val Ser Gln His .Leu Ile Asn Phe Tyr
65 70 75 80

Gly Pro Ala Thr Gly Leu Leu Phe Gly Phe Thr Tyr Tyr Leu Tyr Ala Ala
 85 90 95

Ile Asn Asn Thr Phe Asn Leu Asp Trp Lys Pro Tyr Gly Trp Tyr Cys
 100 105 110

Leu Phe Val Thr Ile Asn Thr Ile Pro Ala Ala Ile Leu Ser His Tyr
115 120 125

Ser Asp Ala Leu Asp Asp His Arg Leu Leu Gly Ile Thr Glu Gly Asp
130 135 140

Trp Trp Ala Phe Ile Trp Leu Ala Trp Gly Val Leu Trp Leu Thr Gly
 145 150 155 160

Trp Ile Glu Cys Ala Leu Gly Lys Ser Leu Gly Lys Phe Val Pro Trp
 165 170 175

WO 94/26901

99

Leu Ala Ile Val Glu Gly Val Ile Thr Ala Trp Ile Pro Ala Trp Leu
180 185 190

Leu Phe Ile Gln His Trp Ser
195

100

CLAIMS

1. Immunogenic composition, capable of inducing antibodies against Helicobacter infection, characterised in that it comprises :

i) at least one sub-unit of a urease structural polypeptide from Helicobacter pylori, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter felis urease, and/or at least one sub-unit of a urease structural polypeptide from Helicobacter felis, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter pylori urease ;

ii) and/or, a Heat Shock protein (HSP), or chaperonin, from Helicobacter, or a fragment of said protein.

2. Immunogenic composition according to claim 1 capable of inducing protective antibodies.

3. Immunogenic composition according to claim 1 characterised in that it includes component (i), which comprises or consists of the Helicobacter felis urease structural polypeptide(s) encoded by the ure A and/or ure B genes of plasmid pILL205 (CNCM I-1355), a polypeptide exhibiting at least 90 % homology with the said polypeptide(s), or a fragment thereof having at least 6 amino-acids and being recognised by antibodies reacting with Helicobacter pylori urease.

4. Immunogenic composition according to claim 1, characterised in that it includes component ii) which is a HSP from Helicobacter pylori, or a fragment thereof.

5. Immunogenic composition according to any of preceding claims characterised in that the HSP is HSP A and/or HSP B encoded by the hsp A and/or hsp B genes respectively, of plasmid pILL689 (CNCM I-1356), or a

101

polypeptide exhibiting at least 75 % homology with the said HSP's, or a fragment of either or both of these proteins having at least 6 amino-acids.

6. Pharmaceutical composition for use as a vaccine in protecting against Helicobacter infection, particularly against Helicobacter pylori and Helicobacter felis, characterised in that it comprises the immunogenic composition of any of claims 1-5, in combination with physiologically acceptable excipient(s) and possibly adjuvants.

7. Proteinaceous material characterised in that it comprises at least one of the Helicobacter felis polypeptides encoded by the urease gene cluster of the plasmid pILL205 (CNCM I-1355), including the structural and accessory urease polypeptides, or a polypeptide having at least 90 % homology with said polypeptides, or a fragment thereof.

8. Proteinaceous material according to claim 7, characterised in that it consists of or comprises the gene product of ure A and/or ure B as illustrated in figure 3, or a fragment having at least 6 amino-acids, or a variant of these gene products having at least 90 % homology, said fragment and said variant being recognised by antibodies reacting with Helicobacter pylori urease.

9. Proteinaceous material according to claim 7 characterised in that it consists of or comprises the gene product of ure I, as illustrated in figure 9, or a fragment thereof having at least 6 amino-acids, or a variant of the gene product having at least 75 % homology, said fragment and said variant having the capacity to activate the ure A and ure B gene products in the presence of the remaining urease "accessory" gene products.

102

10. Nucleic acid sequence characterised in that it comprises :

- (i) at least one sequence coding for the proteinaceous material of any one of claims 6-9 ;
- or (ii) a sequence complementary to sequence (i) ;
- or (iii) a sequence capable of hybridising to sequences (i) or (ii) under stringent conditions ;
- iv) a fragment of any of sequences (i), (ii) or (iii) comprising at least 10 consecutive nucleotides.

11. Nucleic acid sequence according to claim 9 characterised in that it comprises the sequence of plasmid pILL205 (CNCM I-1355), for example the sequence of Figure 3, in particular that coding for the gene product of ure A and for ure B or the sequence of Figure 9 (Ure I), or a sequence capable of hybridising to these sequences under stringent conditions, or a sequence complementary to these sequences, or a fragment comprising at least 10 consecutive nucleotides of these sequences.

12. Expression vector characterised in that it contains a nucleic acid sequence according to claim 10 or 11.

13. Plasmid pILL205 (CNCM I-1355).

14. Oligonucleotide suitable for use as a primer in a nucleic acid amplification reaction, characterised in that it comprises from 10 to 100 consecutive nucleotides of the sequence of claim 10 or 11.

15. Nucleotide probe characterised in that it comprises a sequence according to any one of claims 9 or 10, with an appropriate labelling means.

16. Prokaryotic or eukaryotic host cell stably transformed by an expression vector according to claim 12 or 13.

17. Proteinaceous material characterised in that it comprises at least one of the Heat Shock Proteins (HSP), or chaperonins, of Helicobacter pylori, or a fragment thereof.

18. Proteinaceous material according to claim 17, characterised in that it comprises or consists of HSP A and/or HSP B, having the amino-acid sequence illustrated in Figure 6, or a polypeptide having at least 75 %, and preferably at least 80 % homology with said polypeptide, or a fragment thereof, comprising at least 6 amino-acids.

19. Proteinaceous material according to claim 18 characterised in that it comprises or consists of the HSP A C-terminal sequence :

G S C C H T G N H D H K H A K E H E A C C H D H K K H
or a fragment comprising at least 6 consecutive amino-acids of this sequence.

20. Nucleic acid sequence characterised in that it comprises :

i) a sequence coding for the proteinaceous material of any one of claims 17 to 19 or of any one of the proteinaceous materials of claims 7 to 9 ;
or ii) a sequence complementary to sequence (i) ;
or iii) a sequence capable of hybridizing to sequence (i) or (ii) under stringent conditions ;
or iv) a fragment of any of sequences (i), (ii) or (iii) comprising at least 10 nucleotides.

21. Nucleic acid sequence according to claim 20 characterised in that it comprises all or part of the sequence of plasmid pILL689 (CNCR I-1356), for example the sequence of figure 6, in particular that coding for HSP A and/or HSP B, or a sequence complementary to this sequence, or a sequence capable of hybridizing to this sequence under stringent conditions, or a fragment thereof.

22. Expression vector characterised in that it contains a nucleic acid sequence according to claim 20 or 21.

23. Plasmid pILL689 (CNCM I-1356).

24. Oligonucleotide suitable for use as a primer in a nucleic acid amplification reaction, characterised in that it comprises from 10 to 100 consecutive nucleotides of the sequence of claim 20 or 21.

25. Nucleotide probe, characterised in that it comprises a sequence according to any one of claims 20 or 21 with an appropriate labelling means.

26. Microorganism, stably transformed by an expression vector according to claim 22 or 23.

27. Monoclonal or polyclonal antibodies or fragments thereof, to the proteinaceous material of any one of claims 8 to 10, characterised in that they are either specific for the Helicobacter felis material, or alternatively, cross-react with the gene products of the urease gene cluster of Helicobacter pylori.

28. Monoclonal or polyclonal antibodies according to claim 27 characterised in that they recognise both the Helicobacter felis ure A and/or ure B gene product, and the Helicobacter pylori ure A and/or ure B gene product.

29. Monoclonal or polyclonal antibodies or fragments thereof, to the proteinaceous material of claims 17 or 18, characterised in that they are either specific for the Helicobacter pylori material or, alternatively, cross-react with GroEL-like proteins or GroES-like proteins from bacteria other than Helicobacter.

30. Monoclonal or polyclonal antibodies according to claim 29 characterised in that they recognise specifically the HSP A C-terminal sequence.

31. Use of the immunogenic composition of claim 1 for the preparation of a vaccine suitable for use in man and animals against Helicobacter infection, particularly against Helicobacter pylori and Helicobacter felis.

32. Use of the antibodies of claims 27 to 30 in a therapeutic composition for treating infection by Helicobacter, in particular Helicobacter pylori, Helicobacter heilmannii and Helicobacter felis in man or animals.

33. Method for the production of a pharmaceutical composition according to claim 6, characterised by culturing a transformed micro-organism according to claim 16, and optionally, also a micro-organism according to claim 26, collecting and purifying the Helicobacter urease polypeptide material and where applicable, also the HSP material, and combining these materials with suitable excipients, adjuvants and, optionally, other additives.

34. Use of nucleotide sequences of any claim 15 or 25 for the in vitro detection in a biological sample, of an infection by Helicobacter, optionally following a gene amplification reaction.

35. Kit for the in vitro detection of Helicobacter infection, characterised in that it comprises :

- a nucleotide probe according to claim 15 or 25 ;
- an appropriate medium for carrying out a hybridisation reaction between the nucleic acid of Helicobacter and the probe ;

106

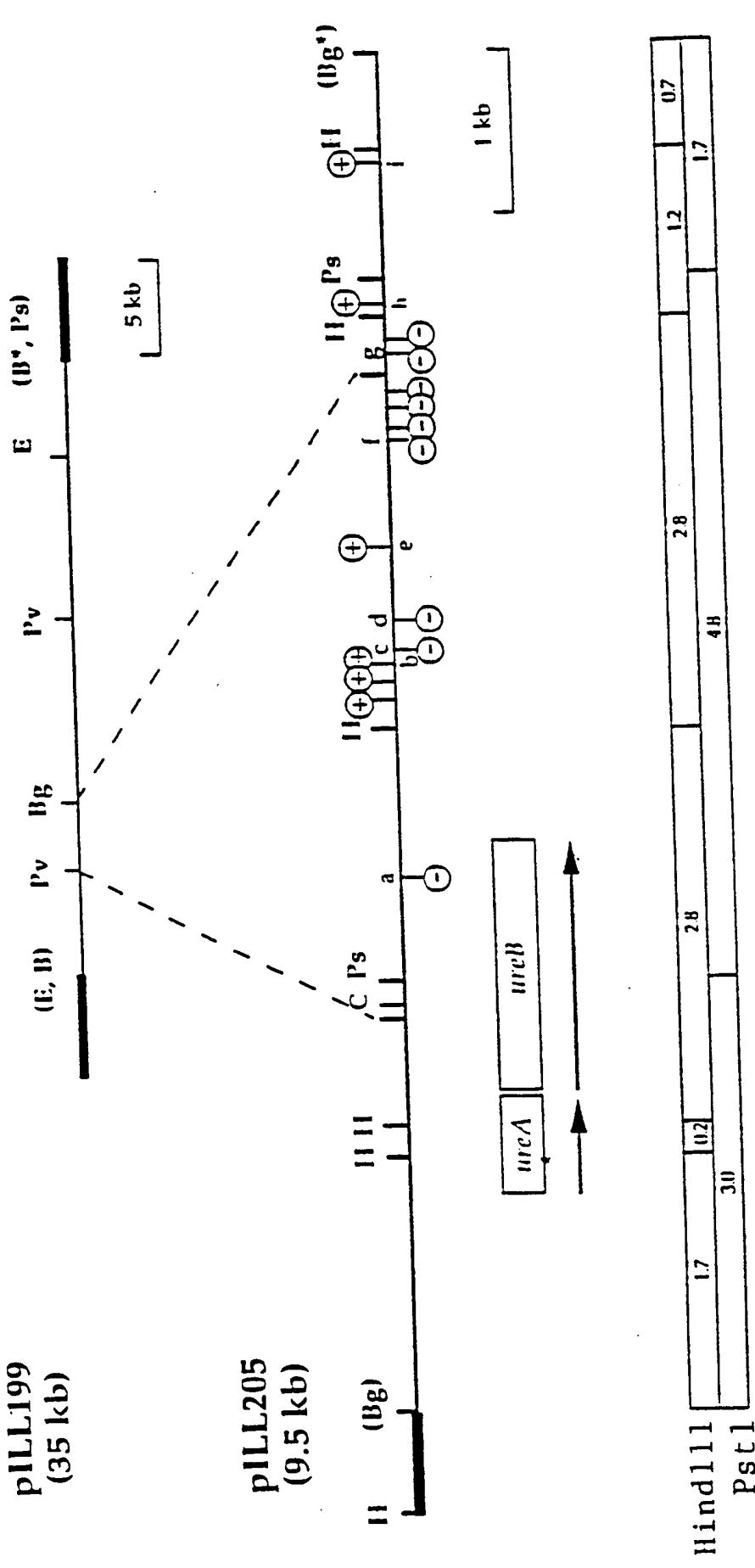
- reagents for the detection of any hybrids formed.

36. Proteinaceous material characterised in that it comprises a fusion or mixed protein including at least one sub-unit of a urease structural polypeptide from Helicobacter pylori or fragment thereof, or from Helicobacter felis or fragment thereof as defined in claims 1 to 3, 5, 7 to 9, and/or a heat shock proteins (HSP) from Helicobacter or fragment thereof, as defined in claims 17 to 20.

37. Purified antibodies or serum obtained by immunisation of an animal with the immunogenic composition according to claims 1 to 5, or with the proteinaceous material or fragment of claims 7 to 9 or 17 to 19, or with the fusion or mixed protein of claims 36.

38. Kit comprising at least the purified antibodies or serum according to claim 37, and optionally, appropriate media or excipients for administration of the antibodies, or labelling or detection means for the antibodies.

1/56



- FIGURE 1 -

SUBSTITUTE SHEET (RULE 26)

1 2 3 4 5 6 7

— 97.4
— 66.2
— 45

— 31

— 21.5

— 14.4

SUBSTITUTE SHEET (RULE 26)

1	2	3	4	5	
					— 97.4
					— 66.2
					— 45
					— 31
					— 21.5
					— 14.4

421/127 ATT AGC TTG AAA GTG AAT AAA GGC GAT CGT CCT GTG CAG GGA TCA CAT TTC CAC
 ile ser leu lys val lys asn lys gly asp arg pro val gln val gly ser his phe his
 481/147 RTC TGC GAA GTG ATT AAG CTC TTG GAC TTC GAT CGC GCA AAA AGC TTT TGC AAA CGC CTA
 phe phe glu val asn lys leu leu asp phe asp arg ala lys ser phe cys lys arg leu
 541/167 GAC ATT GCA TCT GGA ACA GCG GTC CGC TTT GAA CCC GGG GAG GAA AAA AGT GTG GAA CTC
 asp ile ala ser gly thr ala val arg phe glu pro gly glu glu lys ser val glu leu
 601/187 ATT GAC ATC GGC GGG AAT AAG CGC ATC TAT GGC TTT AAT TCT TGC GAT CGC CAA GCC
 ile asp ile gly gly asn lys arg ile tyr gly phe asn ser leu val asp arg gln ala
 661/207 GAT GCC GAT GGT AAA AAA CTC GGC TTA AAA CGC GCT AAA GAA AAA GGT TTT GGG TCT GTA
 asp ala asp gly lys lys leu gly leu lys arg ala lys glu lys gly phe gly ser val
 721/227 AAC TGC GGT TGT GAA GCG ACT AAA GAT AAA CAA TAA GGA AAA ACC ATG AAA AAG ATT TCA
 asn cys gly cys glu ala thr lys asp lys gln OCH ureB Met lys ile ser
 781/6 CGA AAA GAA TAT GTT TCT ATG TAT GGT CCC ACT ACC GGG GAT CGT GTT AGA CTC GGC GAC
 arg lys glu tyr val ser met tyr gly pro thr thr gly asp arg val arg leu gly asp

841/26

ACT GAT TTG ATC TTA GAA GTG GAG CAT GAT TGC ACC ACT TAT GGT GAA GAG ATC AAA TTT
thr asp leu ile leu glu val his asp cys thr thr tyr g1y glu ile lys phe
901/46

GGG GGC GGT AAA ACT ATC CGT GAT GGG ATG AGT CAA ACC AAT AGC CCT AGC TCT TAT GAA
g1y g1y lys thr ile arg asp g1y met ser gln thr asn ser pro ser tyr glu

871/36

ACT GAT TTG ATC TTA GAA GTG GAG CAT GAT TGC ACC ACT TAT GGT GAA GAG ATC AAA TTT
thr asp leu ile leu glu val his asp cys thr thr tyr g1y glu ile lys phe
931/56

961/66

TTA GAT TTG GTG CTC ACT AAC GCC CTC ATT GTG GAC TAT ACG GGC ATT TAC AAA GCC GAC
leu asp leu val leu thr asn ala leu ile val asp tyr thr g1y ile tyr lys ala asp
1021/86

1081/106

ATT GGG ATT AAA GAC GGC AAG ATT GCA GGC ATT GGC AAG GCA GGC AAT AAG GAC ATG CAA ^ 56
ile g1y ile lys asp g1y lys ile ala g1y ile g1y lys ala g1y asn lys asp met gln
1111/116

1141/126

GAT GGC GTA GAT AAT ATT CTT TGC GTA GGT CCT GCT ACA GAG GCT RTG GCA GCT GAG GGC
asp g1y val asp asn asn leu cys val g1y pro ala thr glu ala leu ala glu g1y
1171/136

1201/146

TG ATT GTA ACC GCT GGT GGC ATC GAT ACG CAT ATT CAC TTT ATC TCT CCC CAA CAA ATC
leu ile val thr ala g1y g1y ile asp thr his ile his phe ile ser pro gln gln ile
1231/156

CCT ACT GCT TTT GCC AGC GGG GTT ACA ACC ATT GGA GGA GGC ACA GGA CCT GCG GAT
pro thr ala phe ala ser g1y val thr thr met ile g1y g1y thr g1y pro ala asp

7/56

1261/166	GGC ACG AAT GCG ACC ACC ATC ACT CCC GGA CGC GCT AAT CTA AAA AGT ATG TTG CGT CGA gly thr asn ala thr thr ile thr pro gly arg ala asn leu lys ser met leu arg ala
1321/186	GCC GAA GAA TAC GCC ATG AAT CTA GGC TTT TTG GCT AAG GGG AAT GTG TCT TAC GAA CCC ala glu glu tyr ala met asn leu gly phe leu ala lys gly asn val ser tyr glu pro
1381/206	TCT TTA CGC GAT CAG ATT GAA GCA GGG GCG ATT GGT TTT AAA ATC CAC GAA GAC TGG GGA ser leu arg asp gin ile glu ala gly ala ile 9ly phe lys ile his glu asp trp gly
1441/226	AGC ACA CCT GCA GCT ATT CAC CAC TGC CTC AAT GTC GCC GAT GAA TAC GAT GTG CAA GTG ser thr pro ala ala ile his his cys leu asn val ala asp glu tyr asp val gln val
1501/246	GCT ATC CAC ACC GAT ACC CTT AAC GAG GCG TGT GTA GAA GAC ACC CTA GAG GCG ATT ala ile his thr asp thr leu asn glu ala gly cys val glu asp thr leu glu ala ile
1561/266	GCC GGG CGC ACC ATC CAT ACC TTC CAC ACT GAA GGG GCT GGG GGA CAC GCT CCA GAT ala gly arg thr ile his thr phe his thr glu gly ala gly gly his ala pro asp
1621/286	GTT ATC AAA ATG GCA GGG GAA TTT AAC ATT CTA CCC GCC TCT ACT AAC CCG ACC ATT CCT val ile lys met ala gly glu phe asn ile leu pro ala ser thr asn pro thr ile pro
1681/306	TTC ACC AAA AAC ACT GAA GCC GAG CAC ATG GAC ATG TTA ATG GTG TGC CAC CAC TTG GAT phe thr lys asn thr glu ala glu his met asp met leu met val cys his his leu asp

1771/336
 AAA AGT ATC AAG GAA GAT GTG CAG TTT GCC GAT TCG AGG ATT CGC CCC CAA ACT ATC GCG
 lys ser ile lys glu asp val gln phe ala asp ser arg ile arg pro gln thr ile ala
 1831/356
 1861/366
 GGCT GAA GAC CAA CTC CAT GAC ATG GGG ATC TTT TCT ATC ACC AGC TCC GAC TCT CAG GCT
 ala glu asp gln leu his asp met gly ile phe ser ile thr ser ser asp ser gln ala
 1891/376
 ATG GGA CGC GTA GGC GAG GTG ATC ACA CGC ACT TGG CAG ACA GCA GAC AAA AAC AAA AAA
 met gly arg val gly glu val ile thr arg thr trp gln thr ala asp lys asn lys lys
 1921/386
 GAG TTT GGG CGC TTG AAA GAG GAA AAA GGC GAT AAC GAC AAC TTC CGC ATC AAA CGC TAC -
 glu phe gly arg leu lys glu glu lys gly asp asn asp asn phe arg ile lys arg tyr
 1981/406
 ATC TCT AAA TAC ACC ATC AAC CCC GGG ATC GCG CAT GGG ATT TCT GAC TAT GTG GGC TCT
 ile ser lys tyr thr ile asn pro gly ile ala his gly ile ser asp tyr val gly ser
 2041/426
 GTG GAA GTG GGC AAA TAC GCC GAC CTC GTG CTT TGG AGT CCG GCT TTC TTT GGC ATT AAG
 val glu val gly lys tyr ala asp leu val leu trp ser pro ala phe phe gly ile lys
 2101/446
 CCC AAT ATG ATT ATT AAG GGC GGA TTT ATT GCG CTC TCT CAA ATG GGC GAT GCC AAT GCG
 pro asn met ile ile lys gly gly phe ile ala leu ser gln met gly asp ala asn ala
 2161/466
 TCT ATT CCC ACC CCT CAG CCC GTC TAT TAC CGT GAA ATG TTT GGA ATG CAC CAT GGG AAA AAC
 ser ile pro thr pro gln pro val tyr tyr arg glu met phe gly his his gly lys asn

9/56

2221/486	AAA TTC GAC ACC AAT ATC ACT TTC GTG TCC CAA GCG GCT TAC AAG GCA GGG ATC AAA GAA	2251/496	
lys phe asp thr asn ile thr phe val ser gln ala ala tyr lys ala gly ile lys glu			
2281/506		2311/516	
GAA CTA GGG CTA GAT CGC GCG GCA CCG CCA GTC AAA AAC TGT CGC AAT ATC ACT AAA AAG			
glu leu gly leu asp arg ala ala pro pro val lys asn cys arg asn ile thr lys lys			
2341/526		2371/536	
GAC CTC AAA RTC AAC GAT GTG ACC GCA CAT ATT GAT GTC AAC CCT GAA ACC TAT AAG GTG			
asp leu lys phe asn asp val thr ala his ile asp val asn pro glu thr tyr lys val			
2401/546		2431/556	
AAA GTG GAT GGC AAA GAG GTA ACC TCT AAA GCA GCA GAT GAA TTG AGC CTA GCG CAA CTT			
lys val asp gly lys glu val thr ser lys ala ala asp glu leu ser leu ala gln leu			
2461/566		2491	
TAT AAT TTG RTC TAG GAG GCT AAG GAG GGG GAT AGA GGG GGT TAA TTT AGA GGG GAG TCA			
tyr asn leu phe AMB			
2521		2551	
TTG ATT TAC CTT TGC TAG TTT ATA ATG GAT TTA AGA GAG GTT TTT CGT GTT TTA TAC			
CGC GTT GAA ACC CTC AAA TCT TTA CCA AAA GGA TGG TAA			
2581		2611	

urea

三

二

三

1

H. f. TKLVТИИTPV-----
 H.P. * * * V * * I -----
 P. m. * * * S * * IV 100
 J.B. * * * V*D*ISRENGELOQEALFGSLLPVPSLDKFAETKEDNRI***I
 = = = = = = = =

• • • • • • • • • •
 FLKNEDITI--NAGKEAIISLKVKNKGDRPVQVGSHEFFEVNKL 154
 * * * * * -- E * K * V * V * * V * * * * I * * * * R C * 154
 RVNNALGD*EL* * R*TKTIO*A*H* * * C* * Y* * * Y* * EA* 51
 LCED*CL*L--*I*RK*VI***TS***I***Y***I***PY* 180
 =

- FIGURE 4 (iii) -

unreB

95/ET

- FIGURE 4 (iv) -

PCT/EP94/01625

WO 94/26901

H. f. YKADIGIKDGKIAGIGKAGNKGMDMQDGVDNLCVGPATEALAAEGL
H. P. * * * * * G * * * * K * * S * * * * * G * * *
P. m. V * * * * R * V * * * * P * VQPN * * IVI -- * G * * VV * G * * K
J. b. I * * * * L * S * * * * P * IMN * * FSNMII * AN * * VI * G * * *
==== = === = === = = = = = = =

- FIGURE 4 (vi) -

H. f.
H. P.
P. m.
J. b.

• • •
 LEAIGRTIINTFHTEGAGGGIAPDVIRKMGGEFNILPASTNPTIPE
 MN * * * TM * * * I * * V * * * SV * * P * * * M * Y
 VK * * * VI *
 IR * FK * * RI * * Y * S * * * * * * * * * * * * * * R * I
 =

• • •
 TKNTAEIMDMLMVCIILDKSIKEDEVQFADSRIRPOTIAAEDQLH 351
 * V * T * * 351
 * I * * VD * * I, * * * * * * * * P * * P * * R * * * F * * * E * * * RE * * * I * * 349
 * S * * ID * * I, * KK * * * V * * 622
 =

DNNFRIKRYISKYTINPGIAHGIISDYVGSVEVGKYADLVLWSPAF 441
*****L*****A*****E*****V*****V*****V*****V*****
*****N*****A*****AL*****AH*****I*****K*****L*****I*****D*****
*****R*****A*****A*****Q*****L*****L*****L*****K*****S*****K*****S***** 712
===== ===== ===== ===== ===== ===== ===== =====

FIGURE 6 (ix) -

H. f.	NDVTAAHIDVNTPETYKVVKVVDGKEVTSKADELSLAQLYNL.F	569
H. P.	* * T * * * E * * * * H * F * * * * P * NKV * * * * FSI *	569
P. m.	* NYVP * * ELD * Q * * I * * A * * VPLVCEP * T * * PM * * R * F * *	569
J. b.	* * ALPE * T * D * * S * T * * A * * LLCVSE * TTVP * SRN * F * *	840

ureA	:	74	8	Identity	ureB	:	88	%	Identity
ureA	:	46	8	Identity	ureB	:	62	%	Identity
ureA	:	47	8	Identity	ureB	:	59	%	Identity

PRACTICAL HYDROLOGY

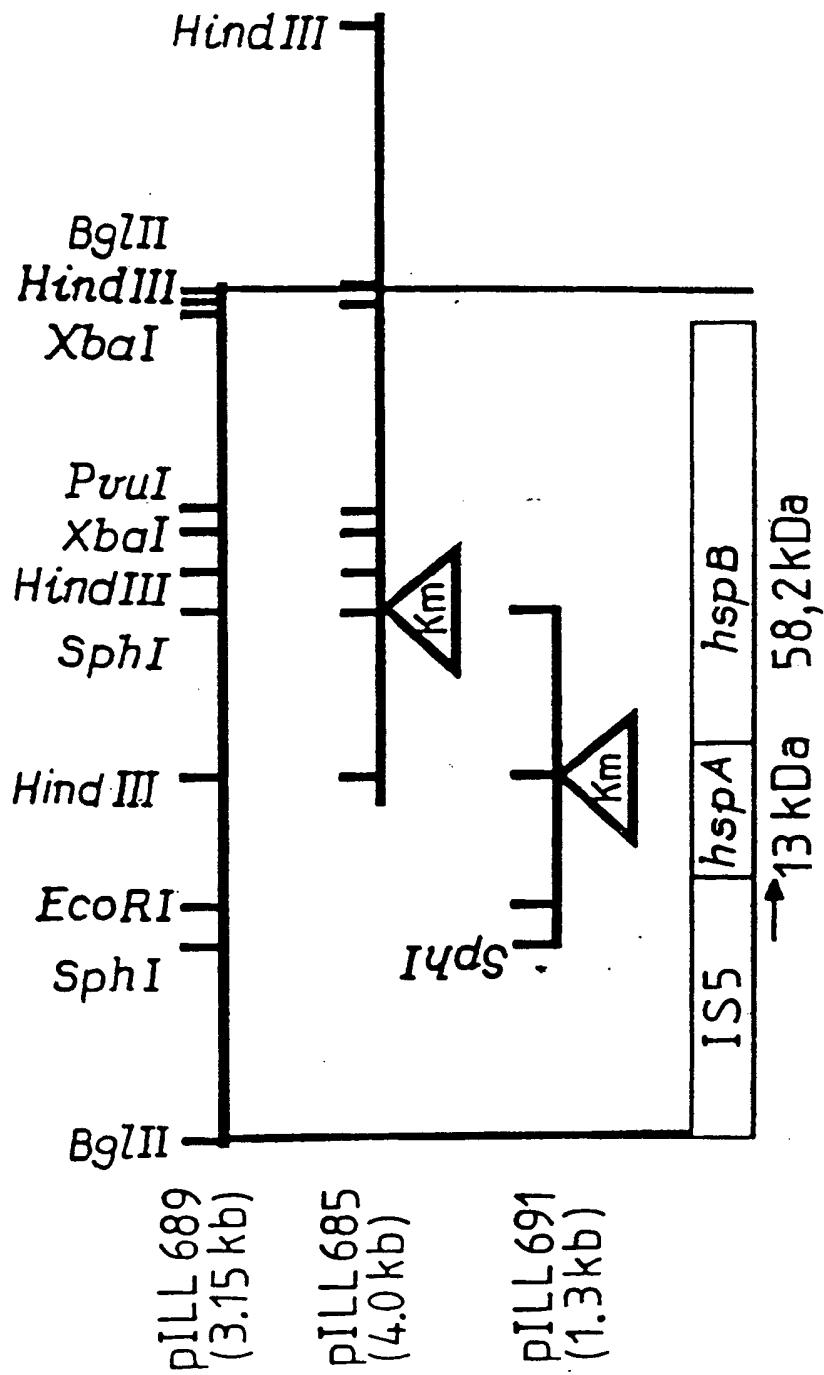


FIGURE 5

21/56

1 ACA AAC ATG ATC TCA TAT CAG GGA CTT GTT CGC ACC CCT AAA AAT GCG CTA TAG TTG	31	
61 TGT CGC TTA AGA ATA CTA AGC GCT AAA TTT CTA TTT TAT TTA TCA AAA CTT AGG AGA ACT	91	
121 GAA ATG AAG TTT CAA CCA TTA GGA GAA AGG GTC TTA GTA GAA AGA CTT GAA GAA GAG AAC met lys phe gln pro leu gly glu arg val glu arg leu glu glu glu glu asn	151/10	
181/21 AAA ACC AGT TCA GGC ATC ATC ATC CCT GAT AAC GCT AAA GAA AAG CCT TTA ATG GGC GTA lys thr ser ser gly ile ile pro asp asn ala lys glu lys pro leu met gly val	211/31	
241/41 GTC AAA GCG GTT AGC CAT AAA ATC AGT GAG GGT TGC AAA TGC GTT AAA GAA GGC GAT GTG val lys ala val ser his lys ile ser glu gly cys lys cys val lys glu gly asp val	271/51	
301/51 ATC GCT TTT GGC AAA TAC AAA GGC GCA GAA ATC GTT ATT GAT GGC GTT GAA TAC ATG GTG ile ala phe gly lys tyr lys ala glu ile val leu asp gly ser cys cys his thr gly asn	331/71	
361/71 CTA GAA CTA GAA GAC ATT CTA GGT ATT GTG GGC TCA GGC TCT TGC TGT CAT ACA GGT AAT leu glu leu glu asp ile leu gly ile val gly ser gly ser cys cys his thr gly asn	391/91	
421/91 CAT GAT CAT AAA CAT GCT AAA GAG CAT GAA GCT CAT GAT CAC AAA AAA CAC TAA his asp his lys his ala lys glu his glu ala cys cys his asp his lys his OCH	451/111	
481 AAA ACA TTA TTA AGG ATA CAA AAT GGC AAA AGA	511	

22/56

479 AAA AAA CAT TAT TAA CGA TAC AAA ATG GCA AAA GAA ATC GAT AGC GCA	509/2 met ala lys glu ile lys phe ser asp ser ala
539/12 AGA AAC CTT TTA TTT GAA CGC GTA AGA CAA CTC CAT GAC GCT GTC AAA GTA ACC ATG GGG	569/22 arg asn leu leu phe glu gln val arg gln leu his asp ala val lys val lys val thr met gln
599/32 CCA AGA GGC AGG AAC GTG TTG ATC CAA AAA AGC TAT GGC GCT CCA AGC ATC ACC AAA GAC	629/42 pro arg gly arg asn val leu ile gln lys ser tyr gln ala pro ser ile thr lys asp
659/52 GGC GTG AGC GTG GCT AAA GAG ATT GAA TTA AGT TGC CCC GTG GCT AAC ATG GGC GCT CAG	689/62 gln val ser val ala lys glu ile glu leu ser cys pro val ala asn met gln ala gln
719/72 CTC GTT AAA GAA GAT GCG AGC AAA ACC GCT GAT GCC GCC GGC GAT GGC ACG ACC ACA GCG	749/82 leu val lys glu asp ala ser lys thr ala asp ala ala gln asp gln thr thr thr ala
779/92 ACC CTG CCT TAT AGC ATT TTT AAA GAG GCG TTG AGG AAT ATC ACC GCT GGG GCT AAC	809/102 thr val leu ala tyr ser ile phe lys glu gln leu arg asn ile thr ala gln ala asn
839/112 CCT ATT GAA GTG AAA CGA CGC ATG GAT AAA GAA GCG CCT GAA GCG ATC ATT AAT GAG CTT AAA	869/122 pro ile glu val lys arg gln met asp lys ala pro glu ala ile ile asn glu leu lys
899/132 AAA GCG AGC AAA AAA GTG GGT AAA GAA GAA ATC ACC CAA GTA GCG ACC ATT TCT GCA	929/142 lys ala ser lys lys val gln gln lys glu ile thr gln val ala thr ile ser ala

959 / 152	AAC TCC GAT CAC AAT ATC GGG AAA CTC ATC GCT GAC GCT ATG GAA AAA GTC GGT AAA GAC	989 / 162	
asn ser asp his asn ile gly lys	ala asp ala met glu lys val gly lys asp		
1019 / 172	GGC GTG ATC ACC GTT GAA GCT AAG GGC ATT GAA GAT GAA TTA GAT GTC GTA GAA GGC	1049 / 182	
gly val ile thr val	glu glu ala lys gly ile glu asp glu leu asp val val glu gly		
1079 / 192	ATG CAA TTT GAT AGA GGC TAC CTC TCC CCT TAC TTT GTA ACC AAC GCT GAG AAA ATG ACC	1109 / 202	
met gln phe asp arg gly tyr leu ser pro	tyr val thr asn ala glu lys met thr		
1139 / 212	GCT CAA TTG GAT AAC GCT TAC ATC CTT TTA ACG GAT AAA AAA ATC TCT AGC ATG AAA GAC	1169 / 222	
ala gln leu asp asn ala tyr ile	leu leu thr asp lys lys ile ser ser met lys asp		
1199 / 232	ATT CTC CCG CTA CTA GAA AAA ACC ATG AAA GAG GGC AAA CCG CTT TTA ATC ATC GCT GAA	1229 / 242	
ile leu pro leu leu glu lys	thr met lys glu gly lys pro leu leu ile ile ala glu		
1259 / 252	GAC ATT GAG GGC GAA GCT TTA ACG ACT CTA GTG GTG AAT AAA TTA AGA GGA GTG TTG AAT	1289 / 262	
asp ile glu gln glu ala leu	thr thr leu val val asn lys leu arg gly val leu asn		
1319 / 272	ATC GCA GCG GTT AAA GCT CCA GGC TTT GGG GAC AGG AGA AAA GAA ATG CTC AAA GAC ATC	1349 / 282	
ile ala ala val lys ala pro	gly phe gly asp arg arg lys glu met leu lys asp ile		

1379/292	GCT GTT TTA ACC GGC GGT CAA GTC ATT AGC GAA TTG GGC TTA GAA AAC GCT	ala val leu thr gly gln val ile ser glu leu gly leu ser leu glu asn ala	1409/302
1439/312	GAA GTG GAG TTT TTA GGC AAA GCG AAG ATT GTC ATT GAC AAA GAC AAC ACC ACG ATC GTA	glu val glu phe leu gly lys ala lys ile val ile asp lys asp asn thr thr ile val	1469/322
1499/332	GAT GGC AAA GGC CAT AGC CAT GAC GTC AAA GAC AGA GTC GCG CAA ATC AAA ACC CAA ATT	asp gly lys his ser his asp val lys asp arg val ala gln ile lys thr gln ile	1529/342
1559/352	GCA AGC ACG ACA AGC GAT TAC GAC AAA GAA AAA TTG CAA GAA AGA TTG GCC AAA CTC TCT	ala ser thr thr ser asp tyr asp lys glu lys leu gln glu arg leu ala lys leu ser	1589/362
1619/372	GGC GGT GTG GCT GTG ATT AAA GTG GGC GCT GCG AGT GAA GTG GAA ATG AAA GAG AAA AAA	gly gly val ala val ile lys val gly ala ser gln val glu met lys glu lys lys	1649/382
1679/392	GAC CGG GTG GAT GAC GCG TTG AGC GCG ACT AAA GCG GCG GTT GAA GAA ATT GTG ATT	asp arg val asp asp ala leu ser ala thr lys ala ala val gln glu gln ile val ile	1709/402
1739/412	GGG GGC GGT GCG GCC CTC ATT CGC GCG GCC CTC ATT TTA CAC GAT GAT	qiy qiy gly ala ala ile leu ala arg ala ala gln lys val his leu asn his asp asp	1769/422

SUBSTITUTE SHEET (RULE 26)

$$= E(\text{efficiency}) -$$

25/56

1799/432 GAA AAA GTC GGC TAT GAA ATC ATC ATG CGC GCC ATT AAA GCC CCA TTA GCT CAA ATC GCT
glu lys val gly tyr glu ile ile met arg ala ile lys ala pro leu ala glu ile ala

1859/452 ATC AAT GCC GGT TAT GAT GGC GGT GTG GTC GTG AAT GAA GAA AAA CAC GAA GGG CAT
ile asn ala gly tyr asp gly val val val val asn glu val lys his glu gly his

1919/472 TTT GGT TTT AAC GCT AGC AAT GGC AAG TAT GTC GAC ATG TTG AAA GAA GGC ATT ATT GAC
phe gly phe asn ala ser asn gly lys tyr val asp met phe lys glu gly ile ile asp

1949/482 2009/502 CCC TTA AAA GTA GAA AGG ATC GCT TTA CAA AAT GCG GTT TCG GTT TCA AGC CTG CCT TTA
pro leu lys val glu arg ile ala leu glu asn ala val ser val ser ser leu leu leu

2039/512 ACC ACA GAA GCC ACC GTG CAT GAA ATC AAA GAA GAA AAA GCG GCC CCA GCA ATG CCT GAT
thr thr glu ala thr val his glu ile lys glu glu lys ala ala pro ala met pro asp

2069/522 2129/542 ATG GGT GGC ATG GGA ATG GGA GGC ATG ATG ATG TAA GCC CCC T'IG C'TG C'TG T'TG
met gly gly met gly gly met gly met gly met met och

2189 GGT ATC ATC TGC TTT TAA ATT CCA TCT TCT RGA ATC CCC CCT TCT AAA ATC CCT TTT T'IG
2219 GGG GGT GCT TTT GGT TTG ATA AAA CCG CTC GCT TTT AAA AAC GCG CAA AAA ACT C'TG

2279 TTA AGC

SUBSTITUTE SHEET (RULE 26)

MAKEIKFSDSARNLILFEGVRLHDIAVKVUTMGPGRNVLIQKSYG
LR*G*D**LQMLA**NA**AQ***VLE***
MA**DV**GND**VKMLR**NV*A***L***KVLD***F*
MA**N**YNED**KIIHK**KT*AE***L***K***H*V*D***F*
TYDEE**RG*ER**LNS**A***L***K***VLE**KW**
Y**DV-*GAD**ALMLQ**DL*A***A***K***T*I*EQ*W*

APSIITKDGVSVAKEIELSCPVANMGAQLIVKEDASKTADAGDG
TV**FEHREM*****M***V***S*T***
T*R***EDKFE***M***V***AN***
S*QV***T***EDKHE***M***V***K***
TN***I***ED*YEKI***F***V***K***D***V***
S*KV***T***S*D*KDKYK*I***K***QDV***NEE***

TYTATVILAYSIFKEGLRNIITAGANPIEVKRGMDKAPEAIIINELKK
 * * * * R * * LV * * I)KAVA * * M * * MDL * * I * * V L * VTKK * QA
 * * * * Q A * IT * * KAVA * * M * * MDL * * I * * VT * AVE * * A
 * * * * EA * YS * * V * * MLD * * I * * VKVVVD * I * *
 * * * * QALV * * V A * * L G L * * IE * * VDKVTET * L *
 * * * * R * * A * * FEK * SK * * V EIR * * V D * V * L * VD * V * A * * *

ASKKVGGKEEITQVATISANSDIHNINGKLIADAMEKVGKDGVIT
 M * * PCKDSKA * A * * G * * * * EA * * AI * * E * * * * E * * * *
 L * VPCSDSKA * A * * G * * * * ETV * * * * E * * D * * * * E * * * *
 I * * P * QHHK * * A * * * * N * AE * * N * * E * * * * N * S * *
 DA * E * ET * * Q * AAT * A * * -G * QS * * D * * E * * D * * NE * * * *
 Q * * P * TTP * * A * * * * G * KE * * NI * SD * * K * * RK * * *

VEEAKGIEDELDVVEGMQFDRGYLSPYFVTNAEKMTAQLDNAYIL
DGN^L^N^Y^I***IN^QQN^SCE^EHPF^*
DG'T^LQ^I***INKE^TGAVE^ESPF^*
F^TV^N^N^N^S^P^TQECSV^EE^LV^*
SNTFGLQ^ELT^R^K*I^GD***RQE^V^EEP^*
KDG^TLN^EIIK***I***INTSKGQKCEFQD^*V^*

LTDKKISSMKDILPPLLEKTMKEGKPLLIIDIEGEALTLVV
*V***V**IREM**SV**GVA*S*R**I**A**
*A***NIREM**V**AVA*A*****V*****A**
IY***GI***F***V**QVVAES*S*R*****A***
*VSS*V*TV**L*****VIOA**S*****V*****S***
*SE***IQS*V*A**IANLVLNIR*KVGLQVVAVK*PGF**L

29/56

NKLRGVILNIAAVKAPGFGDRRKEMI.KDIAVLTGGQVISEEGLISL.
*NM**IVKV C*****A*****Q*****I*****K*****I*****K*****
*TI**IVKV*****A*****Q*****I*****T*****I*****ME*****
*R**AGFRVC*****A*****E*****I*****L*****MK*****
*I***TFKSU*****A*****Q*****M*****A*****V*****T*****
*R*KVG*QVV*V*****N*****NQ*K*M*IA*****A*FG*****GLTLN

ENAEVEF-LGKAKI-VIDKDNTTIVDGKGHSUDVKDRVAQIKT
*G*TL*D-*S**RI*VT*E****I**E*KATEINA*I***RA
*K*TL*D-*Q**RV**N**T**I**V**EEAAIQG****RQ
*TTLAM-****KVIVS*ED***E*L*SKE*IES*CES***K
*TDLSL-****RKV*MT***E***E*A*DTDAIAG***R*
LEDVQPHD***VGEVIVT***DAMLLK*K*DKAQIEK*IQE*IE

- FIGURE 7 A (iv) -

30/56

Q I A S T T S D Y D K E K L Q E R L A K L S G G V A V I K V G A A S E V E M K E K K D
 * M E E * * * * R * * * * V * * A * * * * T * * * * T * * * * A
 * * E E A * * * * R * * * * V * * A * * * * T * * * * A
 * * E D S * * * * R * * * * V * * A * * * * T * * * * R * * * * T * I * * * * *
 E * E N S D * * * * R * * * * A * * * * A * * * * T * * * L * * R * H
 * L D V * * E * E K * * N * * * * S D * * * * L * * G T * D * * V N * * * *

R V D D A L S A T K A V E E G I V I G G G A A L I R A A Q K V H - - - L N - L H D D E K
 * * E * * * H * * R * * * * A * * * * V * * * * V * * * * Q K A L D S - - * K G D N * * Q N
 * * E * * * H * * R * * * * V * * * * S * L A D - - * R G Q N E * Q N
 * * * * * Q H A * L * * * * L P * * T * * V * C I P T L E A F I P I L T N E * * Q
 * I E * * V R N A * * * * A * * * * V T * L Q * * P A L D K - - * K - * T G * * A
 * * T * * N * * R * * * * L * * * C * * L * * C I P A L D S - - * T P A N E * Q *

VGYEIIIMRAIKAPLAQIAINAGYDGGVVVNEVEKHEGHFGFNA
M*IN*LR***ES*MR**VT***EAS***K*AE*KDNY***
*IKVAL***ME***R***VL*C*EEPS***A*T*KGKD*NY*Y*
I*AR*VLK*LS***K***A***KE*AIICQQ*LSRSSSE*YD*
T*AN*VKV*LE***K***F*S*MEF***AEK*RNLSVGH*L*
I*I***K*TL*I*AMT***K***V***SLI*EKIMQSSSEVGYD*

SNGKYVDMFKEGIIDPLKVERIALQNAVSVSSLLTTEATVHEIK
AT*E*G***VEM***L***T***T*M***A***A***M***CM*ADLP
ATEE*GN*IDM***L***T***T*S***Y*A***AG*MI***CM*TDL
LRDA*T***IEA***L***T***T*C***ES***A***AG***LIAD*P
AT*EYE*LL*A*VA***V***T*S***A***IAG*F***V*ADKP
MA*DF*N*VEK***T***V*T***LD*A***A***T*A***VV*T***P

E E K A A P A M P D M G G M G G M G M M	HspB <i>Helicobacter pylori</i>
K K E E G V G A G * * *	HtpB <i>Legionella pneumophila</i>
K N D * * - D L G A A * * *	GroEL <i>Escherichia coli</i>
* * * S S S A - * A * P * A * - * D Y	HYPB <i>Chlamydia psittaci</i>
* K T * * * S D P T G G M G G M D F	GroELL <i>Mycobacterium leprae</i>
* * * D - * G * G A * * * * - * - M * * G * F	63 kDa Human mitochondrial protein P1

Identity : 62.7%
60.5%
59.6%
57.4%
42.5%

Comparison of the GroEL-like proteins from various bacteria

	<i>Helicobacter pylori</i>	YKFQPLGERVL
	<i>Mycobacterium leprae</i>	* * EDKI *
35.6%	<i>Legionella pneumophila</i>	* * IR * * HD * * V
33.8%	<i>Thermophilic bacterium</i>	* LK - * * * D * IV
32.2%	<i>Clostridium perfringens</i>	* SIK * * * D * * V
20.3%	<i>Escherichia coli</i>	* NIR * * HD * * I

VERLEENKTSSGIIIPDNAKEKPLMGVVVKAV---SHKI
 *QAG*A*TM*P**LV**ED***QE*T*V**QGPGRWDE
 *R*M***RT*AG**V***S*T***MR*EII***GAGKVLE
 I*VV*T***A***VIL***T***QE*R*V***GAGRVLID
 IK***A*ET*K***VTGT***R*QEAЕ*V***GPGAIVD
 *K*K*V*T*SAG***VLTGS***AA*STR*E*L*GNGRILE

SEGCKC---VKEGDVIAFGKYKGAEIVLDGVEYMVLELE
DGAKRIPVD*S***IVIYS***G*T**KYN*E*T**LI*SAR
NGDVRA---**V***VL***S*T*V*V***K*LV*MRED
NGQRIGRKS-*V**RVI*S**A*T*VKY**K*Y*I*RES
-GKRTEME-*I**KVLYS**A*T*VKFE*E*T*TI*RQD
NGEVKP-LD**VG*I VI*NDGY*VKSEKIDN*EVLIMS*

DILGIVGSGS CHTGNHDHKHAKHEACCHDHKKH
*V*AV*SK
**M*VIEK
***AVIR
AE
SDILAIIVEA

Comparison of the GroES-like proteins from various bacteria

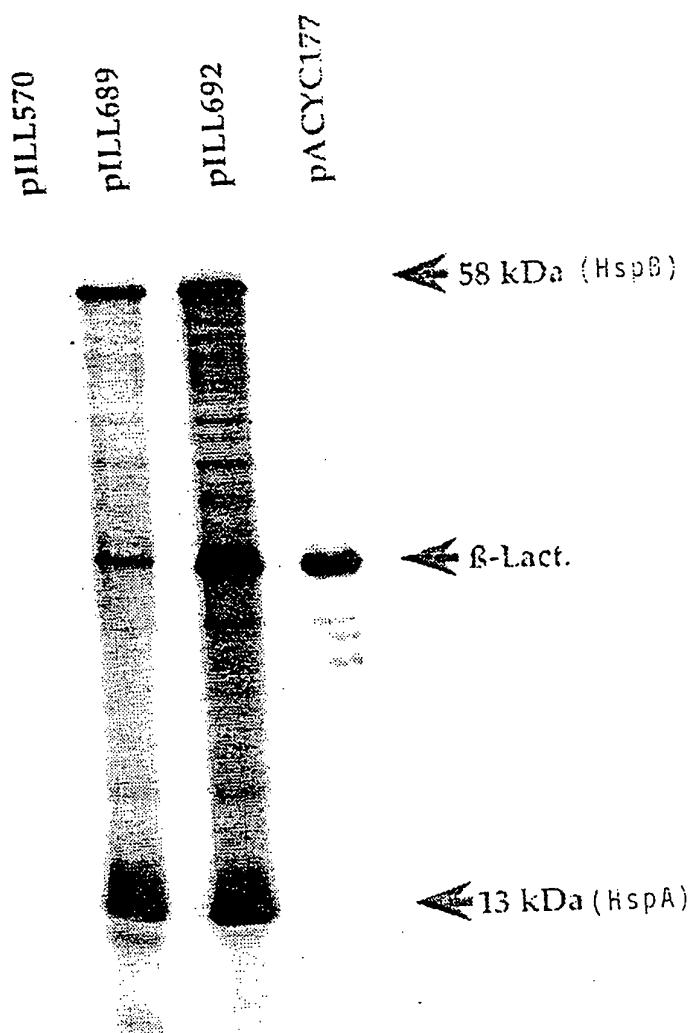


FIGURE 8

1/1
ATG TTA GGT CTT GTG TTA TGT TAT GGT GCG
Met leu gly leu val leu tyr val ala

31/11
GTC GTG CTG' ATC AGC AAC GGA GTT AGT GGG
val val leu ile ser asn gly val ser gly

61 / 21

CTT GCA AAT GTG GAT GCC AAA AGC AAA GCC
leu ala asn val asp ala lys ser lys ala

91 / 31

ATC ATG AAC TAC TTT GTG GGG GGG GAC TCT
ile met asn tyr phe val gly gly asp ser

121 / 41
CCA TCG TGT GTA ATG TGG TCG CTA TCA TCT
pro leu cys val met trp ser leu ser ser

151 / 51
TAT TCC ACT RTC CAC CCC ACC CCC CCT GCA
tyr ser thr phe his pro thr pro pro ala

181 / 61
ACT GGT CCA GAA GAT GTC GCG CAG GTG TCT
thr gly pro glu asp val ala gln val ser

211 / 71
CAA CAC CTC ATT AAC TTC TAT GGT CCA GCG
gln his ile asn phe tyr gly pro ala

241/81
ACT GGT CTA TTG TTT GGT TTT ACC TAC TRG
thr gly leu leu phe gly phe thr tyr leu

271/91
TAT GCT GCC ATC AAC AAC ACT TTC AAT CTC
tyr ala ala ile asn asn thr phe asn leu

41/56

301/101
GAT TGG AAA CCC TAT GGC TGG TAT TGC TTG
asp trp lys pro tyr gly trp tyr cys leu

331/111
TTT GTA ACC ATC AAC ACT ATC CCA GCC GCG
phe val thr ile asn thr ile pro ala ala

42/56

361/121
ATT CTT TCT CAC TAT TCC GAT GCG CTT GAT
ile leu ser his tyr ser asp ala leu asp

391/131
GAT CAC CGC CTC TTA GGA ATC ACT GAG GGC
asp his arg leu leu gly ile thr glu gly

421/141
GAT TGG TGG GCT TTC ATT TGG CTT GCT TGG
asp trp trp ala phe ile trp leu ala trp

451/151
GGT GRT TTG TGG CTC ACT GGT TGG ATT GAA
gly val leu trp leu thr gly trp ile glu

481 / 161
TGC GCA CTT GGT AAG AGT CTA GGT AAA TTT
cys ala leu gly lys ser leu gly lys phe

511 / 171
GTT CCA TGG CTT GCC ATC GTC GAG GGC GTG
val pro trp leu ala ile val glu gly val

541/181
ATC ACC GCT TGG ATT CCT GCT TGG CTA CTC
ile thr ala trp ile pro ala trp leu leu

571/191
TTT ATC CAA CAC TGG TCT TGA
phe ile gln his trp ser OPA

Comparison of the amino acid sequence of the UreI proteins deduced from the nucleotide sequence of the *ureI* gene of *H. felis* and that of *H. pylori*

Percent Similarity : 88.2
Percent Identity : 73.8

First line : *H. felis* Urel
Second line : *H. pylori* Urel

1	KGWMIGI.VII.YVAVVLTISNGVSGLANYDAKSKAIMNYFVGFDSPICVMWS	50
1	: : : : : : : : : : :	
1	...MIGLVLLYVGIVILISNGICGLTKVDPKSTAVMNFFVGGLSIIICNV.V	46
·	·	·
51	LSSYSTFMPPTPPATGPEDVAQVSQULLNFYGPATGLLFGF'TYLYAAINNT	100
51	: . . : : . : : . : : . : : . :	
47	VITYSALNPTAPVEGAEDIAQVSMULTINFYGPATGLLFGF'TYLYAAINNT	96
·	·	·
101	FNLDWKPYGYWYCLFVTINTIPAILSLYSDAILDURLLGITEGDWWAFTIW	150
101	. : : : : : : : : : : : : : :	
97	FGLDWRPYSWYSLFVAINTIPAILSLYSMDLDDIKVLGITEGDWWAFTIW	146
·	·	·
151	LAWGVILWL'RGWIECALGKSLGKFVWPWLAIVEGVITAWIPAWLJFTQIWS	199
151	: : : : : : : : : : : : : :	
147	LAWGVILWL'TAFLENILKIPPLGKETTPWLAJLEGILTAWIPAWLJFTQIWW	195

The Genetic Code

Second Position						
	U	C	A	G		
U	UUU UUC UUA UUG	UCU UCC UCA UCG	UAU UAC UAA* UAG*	Tyr Stop Stop	UGU UGC UGA* UGG	Cys C A Trp
	Phe	Ser				
	Leu					
C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC CAA CAG	His Pro Gln	CGU CGC CGA CGG	U C A G
	Leu					
A	AUU AUC AUA AUG*	ACU ACC ACA ACG	AAU AAC AAA AAG	Asn Thr Lys	AGU AGC AGA AGG	U C A G
	Ile					
	Met					
G	GUU GUC GUA GUG*	GCU GCC GCA GCG	GAU GAC GAA GAG	Asp Ala Glu	GGU GGC GGA GGG	U C A G
	Val					

Abbreviations for amino acids

<i>Amino acid</i>	<i>Three-letter abbreviation</i>	<i>One-letter symbol</i>
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine or glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

49/56

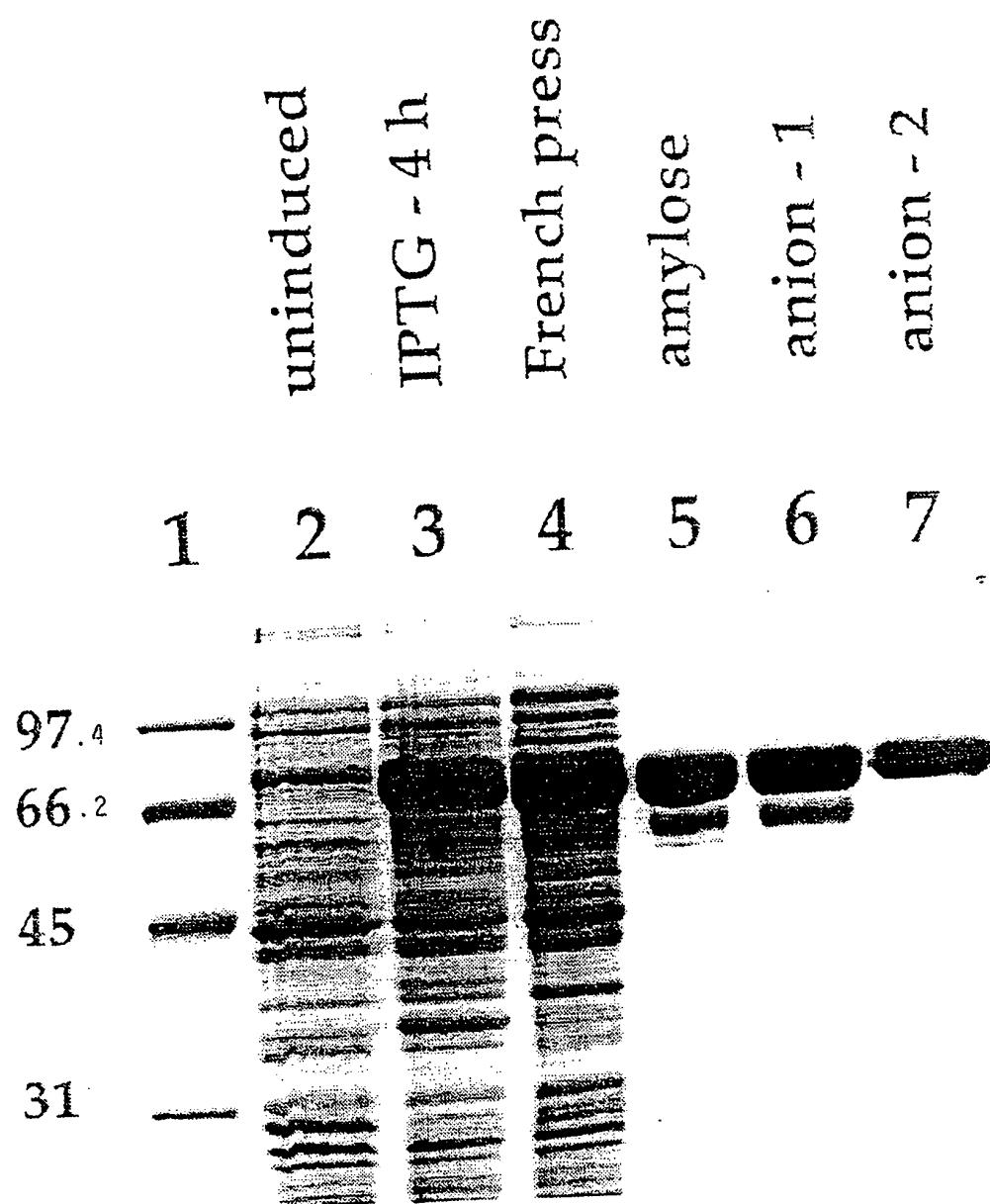


FIGURE 13

50/56

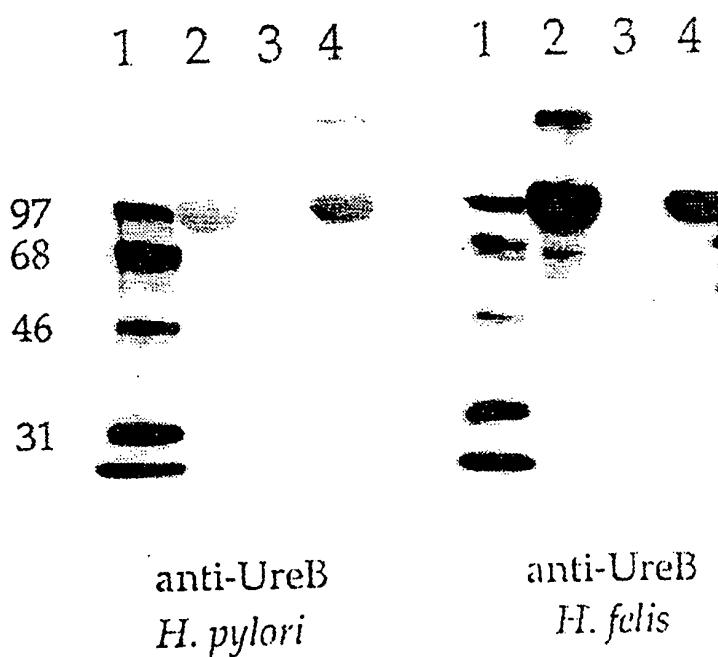
FIGURE 14



SUBSTITUTE SHEET (RULE 26)

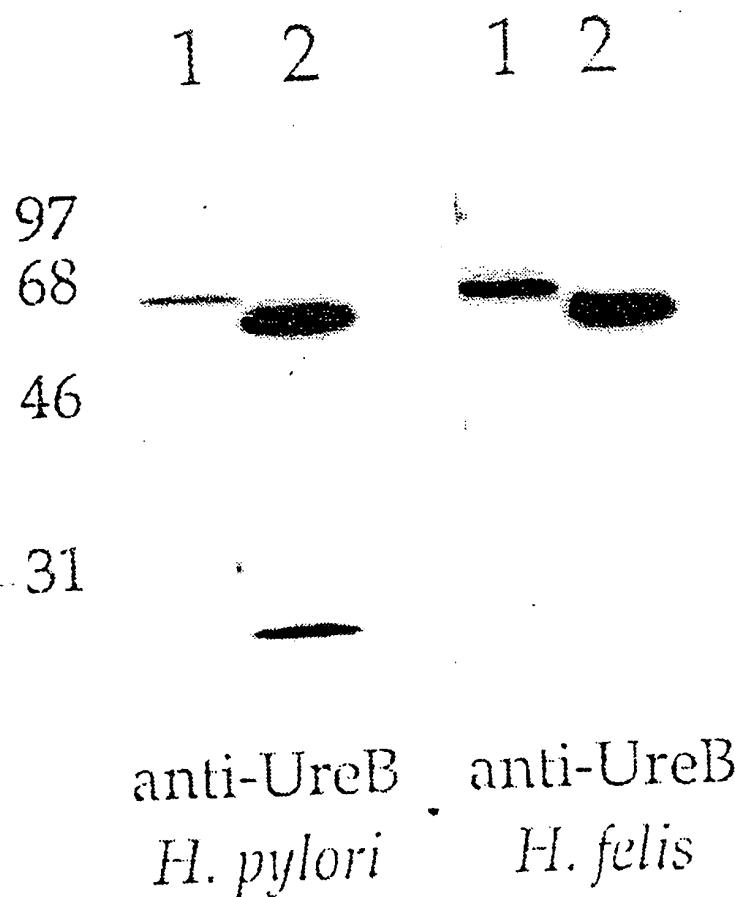
51/56

FIGURE 15

A

54/56

FIGURE 16

B

SUBSTITUTE SHEET (RULE 26)

53/56

FIGURE 17(1)

HSP A - MBP FUSION

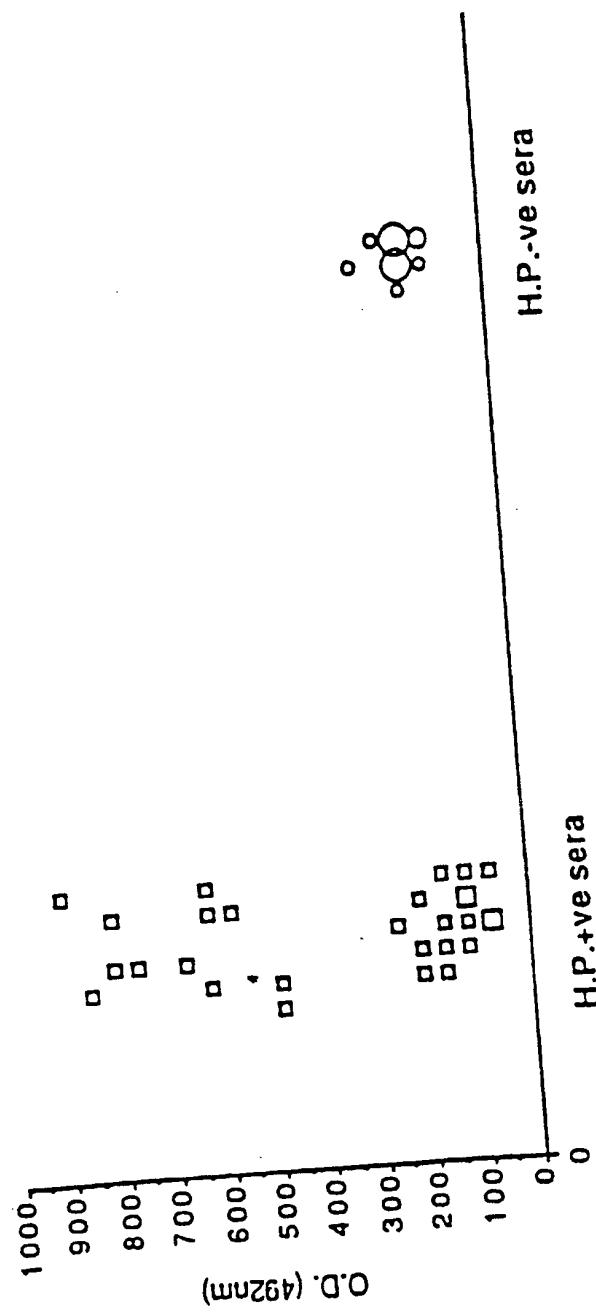
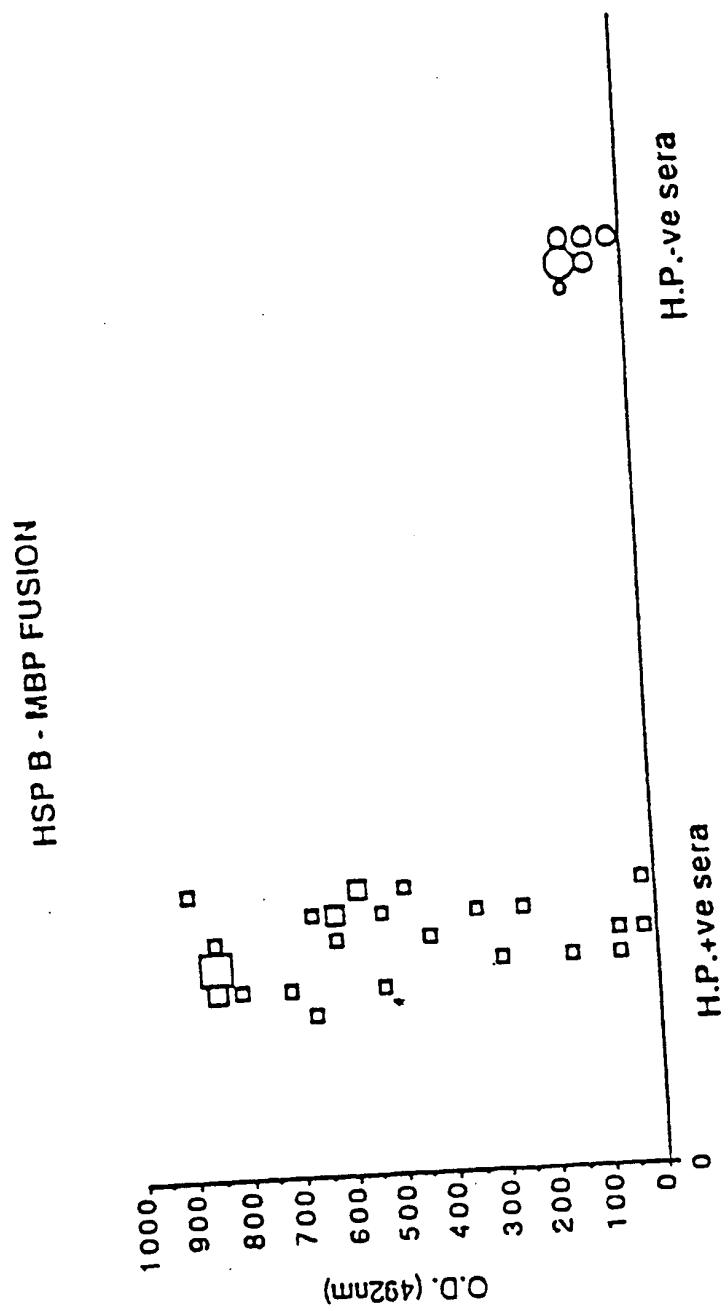
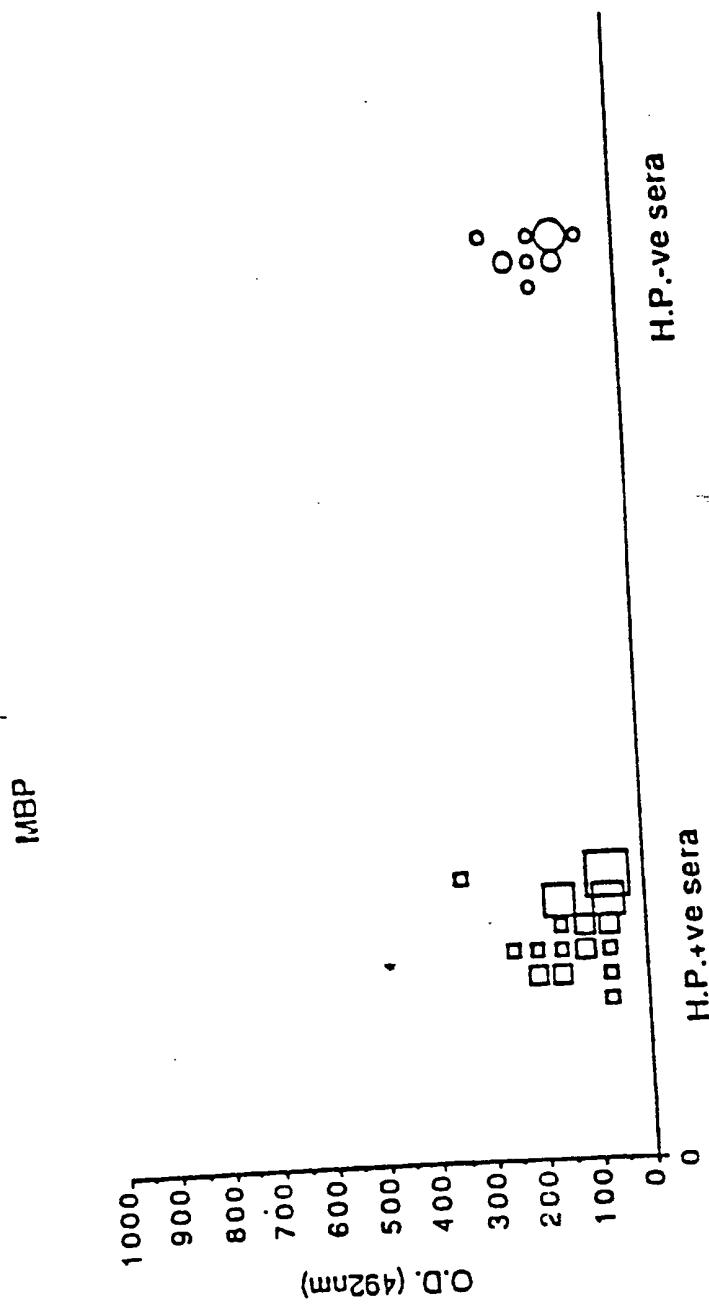


FIGURE 17(i i)



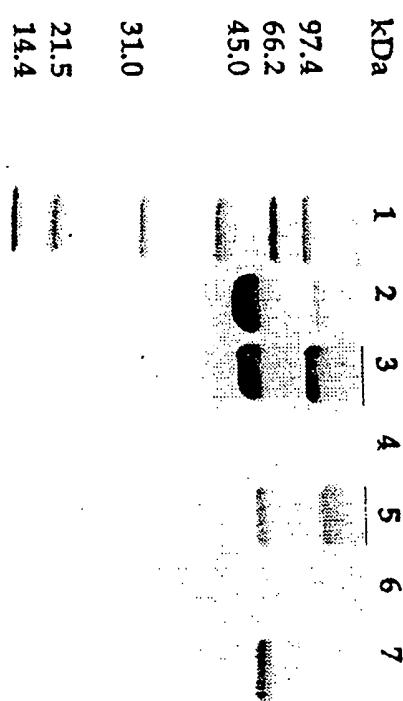
55/56

FIGURE 17(iii)



SUBSTITUTE SHEET (RULE 26)

FIGURE 18



56/56